Summer 8-2021

Matching Equation: Teacher Rates of Praise and Reprimands

Meleah Ackley

MATCHING EQUATION: TEACHER RATES OF PRAISE AND REPRIMANDS

by

Meleah Marissa Ackley

A Thesis
Submitted to the Graduate School,
the College of Education and Human Sciences
and the School of Psychology
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Master of Science

Approved by:

Dr. Brad Dufrene, Committee Chair
Dr. Evan Dart
Dr. Joe Olmi

August 2021

ABSTRACT

Previous matching equation literature has demonstrated variability in student behaviors matching onto available reinforcement rates. While some studies have found that student on-task behaviors matched contingent teacher attention around half the observations (Martens et al., 1990), other studies have found that first grade students' on-task behavior matched contingent teacher attention more than half of observations (Shriver & Kramer, 1997). However, no studies in the current literature have used teacher behaviors as the primary dependent variables (i.e., B1 and B2 in the matching equation). The current study sought to extend the Generalized Matching Equation (GME) further into the classroom, given the concurrent schedules available for teacher behavior. This study examined the effects of a contrived contingency to alter the rates of teachers' praise and reprimand statements. Further, this study assessed the extent of biases and sensitivity to available reinforcement. Although there was variability in biases, sensitivity to reinforcement, and variance explained by the GME, one teacher's behaviors did not indicate bias outside of the contrived contingency in the balanced phase of the study. These results demonstrate that multiple schedules of reinforcement available within the classroom propose challenges to the GME.

ACKNOWLEDGMENTS

I would like to thank Drs. Brad Dufrene, Evan Dart, Jim Moore, Joe Olmi, and Keith Radley for their assistance. Their guidance in this project and throughout my graduate career has been unparalleled. I am also grateful to Kenny Ross, Jennifer Tannehill, and Lauren Layman for collecting data in such an intensive project. I am also thankful to the teachers and schools who volunteered to participate in this study to bridge gaps between basic and applied research.

DEDICATION

I am fortunate enough to have a strong support group of people in my life and throughout this process. To Dustin, for all the moments you have lifted me up and all the sacrifices you have made for me to support my dream of obtaining a Ph.D. Thank you for motivating me to keep going. To my mom, for raising me to be headstrong and loving. These traits have guided my values of helping others and doing so in a way that is valid for them as individuals. To my dad, for always keeping it real and keeping me grounded.

Thank you to all my friends who have assisted me in maintaining a semblance of work/life balance-Taylor Altenberger, Lynda Hayes, Parker Lundy, Bb Mangum-Dear, Andrew Rozsa III, and Jenni Williams. I am so grateful for all the new friends I have made in my graduate career as well. I love you all so much.

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER I - INTRODUCTION

The generalized matching equation states that proportional responding to two alternatives in a free operant paradigm will match the proportion of reinforcement earned on either option (Baum, 1974b; Baum, 1979). Traditionally, the matching equation has been studied by experimental behavior analysts to explain choice behavior in non-humans (Baum, 1974; Crowley & Donahoe, 2004; Herrnstein, 1970; Herrnstein & Loveland, 1975; Herrnstein, 1961) and humans in applied settings (Borrero et al., 2010; Borrero & Vollmer, 2002; Shriver & Kramer, 1997). To date, only a few studies have attempted to apply the matching equation in classroom situations, including investigations involving teacher behavior (Martens, Halperin, Rummel, & Kilpatrick, 1990).

In classroom settings, teachers are faced with a multitude of concurrent operants that compete for their responses. Although an active classroom may seem far removed from laboratory matching equation environments, focusing on a single pair of teacher responses may provide a parallel. If one conceptualizes student appropriate behavior as a potential reinforcer for teacher behavior, there are essentially two common teacher behaviors that result in student behavior. That is, teachers can choose to respond to students' problem behaviors with reprimands (i.e., attempting to punish inappropriate behavior) or praise students' appropriate behavior (i.e., attempting to reinforce appropriate behavior). Similar to basic matching equation studies, reprimands and praise statements can be viewed as concurrent operants with different topographies and capable of being executed independently – just like pigeons' responses to two simultaneously available keys (Herrnstein, 1970). Teacher behaviors (i.e., praise and reprimands) both

presumably result in reduced student disruptive behavior. Yet, researchers have shown that rates of praise and reprimands are typically unequal (Lannie & McCurdy, 2007).

Thus, the present study aimed to expand the matching equation literature by analyzing teacher praise and reprimand behavior through the lens of the generalized matching equation. Additionally, the Good Behavior Game (GBG) provided a standardized context for teacher behavior. The GBG is an interdependent group contingency often applied in classroom settings since its introduction in 1969 (Barrish, Saunders, & Wolf). The GBG has been empirically validated as time-efficient, appropriate, and adaptable in utilizing team competition to increase student on-task behaviors, decrease problem behaviors, and increase teacher praise (Tingstrom, Sterling-Turner, & Wilczynski, 2006). Typically, the GBG operates in one of two ways. First, it may differentially reinforce low rates of behavior (DRL), which are tracked through points earned by reprimands for rule violations. Second, it has also been used to differentially reinforce alternative behaviors (DRA) by tracking points earned by displaying appropriate behavior. This study trained teachers to implement a version of the GBG to concurrently award points contingent on students' appropriate behavior (i.e., deliver praise; DRA) and subtract points contingent on students' disruptive behavior (i.e., provide a reprimand; response cost). Then, with a contrived contingency, this study investigated whether these teacher behaviors conform and are sensitive to the generalized matching equation.

Generalized Matching Equation

To account for deviations from Herrnstein's SME, Baum (1974a) proposed a form of the matching equation known as the Generalized Matching Equation (GME). The

GME is algebraically defined as log(B1/B2) = *a* log(R1/R2) + log *b*. B1 and B2 represent the frequency of each response alternative, and R1 and R2 are the relative rates of reinforcement received from each alternative. *a* represents the slope and sensitivity to relative reinforcement rates. At the same time, *b* reflects the bias of an alternative over the other (Borrero et al., 2010). The GME has more potential for describing human behavior than SME because it accounts for those variations of strict matching by integrating sensitivity and bias into the equation (Greguson, 2008).

The additional parts of the equation also help determine if undermatching, overmatching, or bias occurs (Baum, 1974, 1979). Undermatching occurs when the log ratio of responding is increased less than one unit while delivered by a one-unit increase in log ratio of reinforcement (Borrero et al., 2010). According to Baum (1974b), the rate of responding is less than predicted by matching. On the other hand, in overmatching, the rate of responding is more than expected by the matching equation. In bias, if *b* is greater than the value for B1 in the GME formula, there is a bias for B1. If *b* is negative, there is a bias for B2. Deviations from bias can be explained by response bias (e.g., organism's preference, response effort), inconsistencies between response reinforcement and scheduled reinforcement, the variance of potential reinforcers (e.g., praise vs. social disapproval), and variance of schedules (e.g., VI vs. VR).

Matching Equation and Human Behavior

Although the matching equation is most known from animal research (Baum, 1974a; Baum, 1974b; Herrnstein, 1961; Herrnstein, 1970), matching analyses have extended to human behaviors. McDowell (1988) utilized the matching equation in one of the first cases to apply the matching equation to a natural human environment with self-

injurious behaviors. Herrnstein's single-alternative matching equation (i.e., quantitative law of effect; 1970) described the rate of self-injurious behaviors matching the rate of verbal reprimands from family members as they all watched television. A reversal design concluded that the self-injurious behaviors were reinforced by the family's verbal reprimands/attention. Although other responses were available for reinforcement in that type of environment (i.e., the television), the matching equation still accurately described the client's behavior in relation to others' behaviors (McDowell, 1988).

Reed, Critchfield, and Martens (2006) extended the generalized matching equation onto the football field. They wanted to know if offensive outcomes could predict the play calling for the National Football League (NFL) in 2004. They used descriptive data from websites reporting football statistics to determine the relative ratio of passing versus rushing plays called by head coaches across 32 NFL teams. The ratios of passing and rushing plays were compared to the relative ratio of reinforcement earned by each type of play. The relative ratio of reinforcement was defined as the number of yards gained. Except for a few individual teams, the generalized matching equation explained head coaches' preference for either type of play (i.e., running or passing plays). Undermatching explained most play calling, along with a bias for rushing plays to be called.

Similarly, Vollmer and Bourret (2000) found that the generalized matching equation explained male and female college basketball players' preference for two-point and three-point shots. The subjects were 13 males and 13 females who played for a large National Collegiate Athletic Association (NCAA) Division I school. This was the first study that accounted for potential reinforcer amounts for humans in the matching

equation. Three-point shots were equated as 1.5 times more valuable than two-point shots. Concurrent choices were available to each player because he/she could choose to attempt a two-point or a three-point shot. Results showed that male and female players who attempted and scored from a three-point range were more likely to attempt three-point shots than males and females who attempted and often failed to score from a three-point range. Thus, the players who had more playing time and attempted more shots were where the matching equation predicted shot distribution.

Matching Equation in the Classroom

Since reinforcement is often under the teacher's control, Martens et al. (1990) evaluated Herrnstein's 1970 matching theory with contingent teacher attention with a 6-year-old boy's behavior in a summer school program. It was the first administration of Herrnstein's equation related to classroom behavior and contingent teacher attention. For two weeks, teacher attention, on-task behaviors, and off-task behaviors were observed for a total of 322 minutes. The duration of contingent teacher attention was applied as an approximation of accessible reinforcement. On-task behaviors adhered to the SME for an average of 51% of the observations, whereas off-task variance matched the rate of reinforcement an average of 47% of observations. Thus, on average, the participant's behavior only matched reinforcement for teacher behavior about half of the observation time, which could be a natural occurrence of a classroom environment with multiple concurrent operants. Additionally, the matching equation used (i.e., Herrnstein's hyperbola) did not account for 49% of the choice behavior variability. This may have also been due to the difficulty in determining all of the concurrently available reinforcement. In Herrnstein's hyperbola, choice behavior is only based on relative rates

of reinforcement, but extraneous variables likely affect choice behavior, along with immediacy and sensitivity to reinforcement. Thus enters the GME to account for such variables outside of relative reinforcement rates.

Shriver and Kramer (1997) studied the GME regarding student behavior in a first grade and fourth grade classroom. Teacher behavior (e.g., listening, business management, instruction) was analyzed for 30 seconds after student behaviors (e.g., listening, waiting, task appropriate) as a measure of reinforcement for student behavior. For the first graders, 73.5% of their behaviors (for a median variance) adhered to the GME. Undermatching ($a < .90$; Baum, 1979) also occurred in their respective generalized matching equations, such that the behaviors alternated between each other less than predicted by the GME. Since the mean biases were not statistically significant, there was no observable bias for a specific behavior or a reinforcement schedule. However, for 2 out of 4 fourth grade students' behaviors, teacher behavior did not match, thus there may have been other extraneous variables unaccounted for that competed for students' reinforcement rates. Since there are competing contingencies in an applied classroom setting, it is helpful to choose an intervention like the GBG to control for such contingencies by providing the same classroom management technique to teachers within the study.

The Good Behavior Game

The Good Behavior Game (GBG) has effectively decreased not only disruptive behaviors but also increased academic and prosocial behaviors in several settings, including elementary (Barrish et al., 1969), preschool (Swiezy, Matson, & Box, 1993), and high school (Ford, 2015) classrooms. Despite the extensive literature on the efficacy

of the GBG, there is little research comparing the effects of GBG rule following (Swiezy et al., 1993) versus rule violations (Barrish et al., 1969) on student and teacher behavior. It is important to note that awarding team points for rule following can be described as positive reinforcement, while awarding points for rule violations defines a response cost, which is a form of negative punishment. It is also arguable that awarding points for rule violations could be viewed as positive punishment. When responding to rule following, teachers provide reinforcement through praise (e.g., with a token), and rule violations are disregarded. Conversely, when teachers respond to rule violations, the team loses a point when rules are not followed (Tanol, Johnson, McComas, & Cote, 2010).

Variations of the Game

In a Belgian elementary classroom, a variation of the GBG was applied to reduce disruptive behavior (Leflot, Lier, Onghena, & Colpin, 2010). Students were followed from their second-grade classrooms through the third grade. Before the game was employed, the randomized control and experimental groups exhibited the same levels of disruptive behavior. Classrooms were divided into teams of 4-5 students. When rules were followed, the team received praise, but the team would lose 1 of their five cards (symbolizing rewards) contingent upon rule violations. If the team had at least one card at the end of the game, they received a prize. Initially, the game was implemented at 10-minute intervals. This was gradually lengthened until the game persisted for half of the school day. At the end of the first-grade year, academic engagement had significantly improved, while disruptive behavior decreased considerably. The participants were followed into third grade, but the results failed to generalize over time. Subsequently,

after the re-implementation of the GBG, there were similar results to those found at the end of the second grade year.

Before GBG implementation, there were no significant differences between second and third grade teachers' use of praise or negative remarks. The second grade GBG teachers used less negative attention and notably more praise, particularly at the end of the school year. The third grade GBG teachers employed more praise as well. This could result from the GBG children (compared to the control group) already exhibiting lower frequencies of disruptive behavior at the end of the second-grade school year. Although Leflot et al. (2010) theorize that declines in teachers' negative management behaviors are related to decreases in children's disruptive behaviors, this should be confirmed with data. Plus, there are still unexplained differences in teacher behavior that should be addressed, such as teachers' motivating operations for behaviors and if such preferences (i.e., biases) may be manipulated.

Teacher Preference and Acceptability

Teacher preference may tie into motivating operations. Asking a teacher's preference for a classroom management technique may increase a teacher's acceptability of an intervention. This may further increase the acceptability by incorporating their preferences as choices for the types of interventions they may implement with their students. For example, Dart and colleagues (2012) allowed teachers to "test-drive" interventions. This allowed the teachers to test interventions such as self-monitoring, modified Check-in/Check-out, response cost, and behavior specific praise. After testing the interventions, teachers ranked the interventions from most to least acceptable. Then, the authors allowed the teachers to re-implement their most preferred intervention, which

resulted in higher treatment integrity levels for the preferred intervention contrasted with the other interventions. Additionally, students' academic engagement increased during the preferred intervention implementations. The incorporation of teacher preference may have increased the likelihood that teacher acceptability and integrity would increase.

Tingstrom (1994) examined teacher acceptability of GBG-response cost and GBG response cost plus merit strategy. In one experiment, intervention and behavior problem severity were independent variables, with the interventions comprised positive reinforcement, GBG, GBG + M (merit points), and response cost. The dependent variable was the score on the Intervention Rating Profile (IRP-15). In this study, 89 general education and special education teachers were given case descriptions of problematic behaviors in the classroom. They were given the four intervention options to rate acceptability. The GBG was found equally acceptable as positive reinforcement and response cost classroom management techniques. In a second experiment, the GBG was assessed to find potential extraneous variables related to the student's age and the source of the rationale for choosing the intervention options (teacher, psychologist, or no explanation). In this study, there were 115 elementary and middle school regular and special education teachers as participants. They were also given case descriptions, the four intervention options, and the IRP-15. The age of the students nor the types or severity of problem behavior(s) had significant effects on acceptability. It is noted that the study has limited variability because it was an analog study that was not based in a natural classroom setting. Although this study demonstrates that teachers support these interventions, more research is necessary to replicate these effects in the classroom setting.

Tanol et al. (2010) attempted to extend Tingstrom's findings (1994) to a natural setting by implementing a single-case A/B/A/C/B/C reversal design focusing on GBG response cost versus GBG reinforcement. They also studied how teacher attention to both GBG response cost and reinforcement affected teacher and student behavior. The target students were 6 kindergarten students who were identified as the three most disruptive students in each classroom. In baseline, teachers did not implement reinforcement or group contingencies. They continued with their day-to-day routines. Each teacher was randomly assigned either GBG response cost or GBG reinforcement for the first B phase. In both versions, observations were on a 10 second partial interval schedule for ten-minute intervals during carpet time. For Classroom 1, response cost was first introduced. In the return to baseline for rule violations in Classroom 1, the rules were not followed an average of 50% of the time. When GBG reinforcement was introduced as the C phase, rule violations immediately decreased to an average of 25%. For the second phase B condition, the GBG response cost was re-implemented for Classroom 1. At this point, rule violations increased up to 35%. After that phase, the last GBG reinforcement condition was in place. The results were similar to those found in the first GBG reinforcement condition. Classroom 2 had similar results but began with GBG reinforcement.

In the GBG response cost version, the teams started the game with four stars, and removal of the stars was contingent upon rule violations of any team member. When a student did not follow a rule, the teacher responded by saying the rule was not obeyed and praised the other team for abiding by the rules. In GBG reinforcement, all groups began the game with unmarked posters. Teacher attention was only distributed for rule

following. No consequences were earned for rule violations. Both types resulted in a decrease in rule violations from the students, yet GBG reinforcement was more reliable in reducing rule violations, and the teachers preferred it. Teachers reported that reinforcement was preferred because it promotes a more positive classroom environment than response cost environments. One may also argue that by increasing the likelihood of teacher praise statements, the students are reinforced more often, which may relate to teacher reinforcement; however, informal interviews were the basis for teacher preference comparing GBG rule following versus rule violations. A more direct test of preference would have involved allowing teachers to choose which version to run (i.e., a "test drive"), but teachers were not given that choice in this study.

Contingent Teacher and Student Behavior

Although there have been no studies found to date on the GBG and matching equation, a few studies have analyzed the reciprocal effects of student behavior on teacher behavior (Elswick & Casey, 2011; Lannie & McCurdy, 2007). Reciprocal effects are teachers' responses resulting from students' appropriate and problem behaviors. Reciprocal effects refer to the degree to which teachers may modify their responses (i.e., praise and reprimands) to increases and decreases in students' appropriate and inappropriate behaviors. Lannie and McCurdy (2007) studied the effects of the GBG on teacher's praise, neutral, and disapproval statements with student behaviors in an urban school district. Implementation of the GBG increased student on-task behaviors and decreased targeted disruptive behaviors, but the change in student behavior had little effect on teacher behavior.

In an A-B design conducted by Elswick and Casey (2011), they targeted teacher responses to student behaviors to replicate and extend the findings of Lannie and McCurdy (2007). The teacher scanned the room for student rule violations (which resulted in points gained for the teacher) and student positive behaviors (points gained for students). At the end of the week, whichever team (all the students or the teacher) had the most points earned received a reward. Once the GBG was implemented, targeted student behaviors (e.g., talk outs, out of seat, and disrespectful behaviors) decreased. The teacher behavior data indicated an increase in praise statements (from M=3.5 to M=13.2) and a decrease in disapproval statements (from M=7.85 to M=1.5). In other words, when given an option to award points and subtract points, the teacher's praise statements increased, and disapproval statements decreased.

The literature provides few evidence-based interventions that address the effect of student behavior on teacher behavior. These effects are significant to study because previous research has shown that teacher reprimands for inappropriate social behaviors are significantly higher than teacher praise statements for appropriate social behaviors (Beaman & Wheldall, 2000). Coincidentally, student maladaptive behaviors can increase with inappropriate teacher responses (Kodak, Miltenberger, & Romaniuk, 2003). The data show that teachers may not be sensitive to the contingencies at work and that maladaptive behaviors can increase with inappropriate teacher responses. Ultimately, an examination with a matching equation may explain the relationship of teacher praise statements to reprimands when a contrived contingency is implemented.

According to Poling et al. (2011), Herrnstein's 1961 and 1970 articles on matching equations have been the top-cited articles in the *Journal of Applied Behavior*

Analysis and the *Journal of the Experimental Analysis of Behavior*. This further suggests that matching equations could conceivably unite basic and applied research. The purpose of this study was to extend the applied research on matching equations by utilizing the generalized matching equation framework to determine if teachers' responses were consistent with the rate of reinforcement they receive for engaging in disapproval statements and behavior specific praise statements.

Research Questions

The following research questions were evaluated:

1. Does the proportion of teachers' praise statement rates to reprimand rates alter when a contrived contingency is implemented that conforms to the generalized matching equation, or are biases present?
2. Does the proportion of teachers' praise statements to reprimands match, undermatch, or overmatch the contrived contingency when the proportion of reinforcement is modified?

CHAPTER II - METHOD

Participants and Settings

Data were collected in a small urban school in a southeastern state. The school has a student-teacher ratio of 14:1, wherein most students identified as African American (86.7%) and 100% of students qualified for free or reduced lunches. Participants included three licensed elementary school teachers who volunteered for the study to improve their classroom management skills. Teachers A and C both taught second grade, while Teacher B taught first grade. All three teachers identified as the following: female, Caucasian, born between 1994-1989, Bachelor's Degrees as highest levels of education, and less than two years teaching experience (see Appendix A for the demographics form). Before data collection, the primary investigator gained approval to conduct the study through the University of Southern Mississippi's Institutional Review Board (IRB; Appendix B). Afterward, permission was secured by the school district's Special Projects and Curriculum Director. The school's principal provided approval via email. Then, teachers in Classrooms A and B were recruited from teacher training for teachers new to the school district. Initially, another teacher was also recruited from that training, but she became a Teacher's Assistant who was no longer in charge of classroom management. Thus, Classroom C's teacher was recommended by the principal and subsequently volunteered for the study after communication with the first author. After the principal investigator reviewed the study with the teachers, teacher consents were signed by all three teachers (see Appendix C for details). Of note, the teachers were blind to the VI schedules and the order of the phases, but they were told they would be taught the GBG

as a classroom management technique, and this would involve both students and teachers receiving rewards.

Materials

Class-Wide DBR

The teachers filled out a class-wide Direct Behavior Rating (DBR; Appendix D) after each session. DBR was chosen due to research findings that DBR data find similar results to systematic direct observations for class-wide student behaviors (Riley-Tillman, Methe, & Weegar, 2009; Riley-Tillman, Chafouleas, Sassu, Chanese, & Glazer, 2008). Printed directions were provided to the teacher on how to fill out the ratings on the form, and observers were present when the teachers filled out the form in the event teachers had any questions. Observers were instructed not to provide their opinions of students' on-task behavior. Teachers were told verbally and within the printed directions to rate the overall student behavior instead of focusing on particular students. The sheet also included the rating scale from zero to ten, which had descriptors of zero, meaning that students were never on-task or were on task 0% of the time observed, five indicating 50% or "sometimes" on-task, and ten meaning 100% or "always" on-task for the observation period. After 100% of the sessions, teachers filled out the class-wide DBR form. The rating scale was from 0-10 or 0-100% for class-wide on-task behavior.

Teacher Script

The teacher script (Appendix E) increased the likelihood for consistency of the GBG across sessions. It included a review of the student expectations for the game (e.g., the criterion for teams to win; points added for following rules; points taken away for not following rules). It also included reminders for the teacher to state the classroom rules,

deliver praise when adding points, deliver reprimand when removing points, the game duration was 20 minutes, and provide winning students a prize.

Rules Posters

Each teacher had a rules poster displayed during all intervention phases. Rules were based on teacher expectations and developed with each teacher according to students' off-task behaviors in the baseline.

Teacher and Student Preferences

Before interviews with the teacher on preferences for their rewards, the primary investigator completed brief informal interviews with the principals. They approved school supplies for teachers' rewards. Then, all student and teacher rewards were approved by the teacher and primary investigator prior to utilization. Students and teachers approved their respective rewards through informal preference assessment interviews. In the informal student interviews, teachers asked the students what they would work for, and students replied they would work for colorful pencils, chocolate candy, sour candy, suckers, and small bouncy balls. In the informal interview with teachers, the primary investigator provided a list of school options from which the teacher chose 5 items.

Reward Menus

From the chosen items, reward menus were created for each teacher (Appendix F). Rewards were provided by the primary investigator (e.g., pencils, candy for students; Expo markers, post-it notes for teachers). Items reported by the teachers as more preferred cost more points. Items included Expo markers, glue sticks, 2 AA batteries, pen, pencil, sticky note block, or roll of tape. Each teacher had five items on their reward

menu. Item points were based on the average number of opportunities for each teacher to earn rewards (per session for both praises and reprimands). For example, on average, in Classroom A, praise occurred every 37 seconds, and reprimands occurred every 120 seconds in the GBG baseline. So, once accounted for 1200 seconds per observation, 32 opportunities for praise on average, and ten opportunities for reprimands on average, then there are 42 opportunities per session to get rewarded. The primary researcher approximated that there would be seven sessions per intervention phase. So, the primary researcher multiplied three (intervention phases) by seven (sessions per intervention phases), which equaled 21; then multiplied 21 by 42 opportunities, equaling 882. Authors hypothesized that teachers would match the VI schedule 80% of the time, which would mean gaining 705.6 points, since 80% of 88 is 705.6. Based on the point system for Classroom A's reward menu (Appendix G), teacher A could hypothetically earn an expo marker after two sessions. Three phases=21. 42*21=882; 80% of which is 705.6.

Whiteboard

The teacher's use of reprimand or praise statements following each VI phase was recorded on the primary investigator's whiteboard in the back of the classroom, visible to the teacher. When meeting the VI schedule for praise or reprimands, a tally mark was written on the board. Tally marks were not differentiated between points for praises or reprimands.

Data Sheet and Countee App

In addition, the observers had a data collection sheet (Appendix G) to record when the teacher had met a VI schedule. Upon meeting a VI schedule, the teacher received a tally mark on the whiteboard in the back of the classroom. A black tally mark

indicated that the teacher met an available reprimand or praise schedule; they were not differentiated on the observer's whiteboard. At the end of each observation, observers recorded the number of points the teacher had received that day. The number of points received could then be exchanged for something on the reward menu. The teacher could also choose to wait on selecting an item from the reward menu to use points for an item of higher value (e.g., Expo marker). Thus, observers also recorded how many points the teacher utilized for rewards.

To collect the frequency of reprimands, praises, and when the teacher met the VI schedule, the "Countee" app was employed (see Appendix H for details). During each observation, the observer(s) pressed the "start" button to begin the timer for the twenty-minute observation within Countee. Then, they immediately pressed the "praise time," followed by the "rep time" buttons to begin the timers for praises and reprimands. When reinforcement became available for either VI schedule, the observers stopped those timers. When a teacher met a VI schedule, the "praise delivered" or "reprimand delivered" buttons were employed, and a tally was marked on the whiteboard. If praise was delivered and met the VI schedule, the "praise time" button was pressed again. Observers were trained to press the "praise time" buttons according to the VI schedules on their datasheets. Once a VI schedule was met, observers checked it on the data sheets (Appendix G). At any time a teacher delivered praises or reprimands, then the "praise" or "reprimand" buttons were pressed to count each instance of behavior.

Dependent Variables

Praise Statements

Teacher praise included general praise, encompassing non-directed praise statements (e.g., “Good job!”; “Nice work!”; “There you go.”). Teacher praise also involved directed praise statements (e.g., “Good job, class!”; “Team A just gained a point.”). “Nice work, Sally!”) and behavior specific praise statements (e.g., “Good job sitting down, Bobby!”; “I like the way Johnny is raising his hand before speaking”). Praise statements also included instances in which the teacher awarded only points for appropriate behavior within the context of the GBG. Non-examples included any praise or praise points delivered by someone other than the teacher. All types of teacher praise collapsed into a single frequency count of praise statements.

Reprimands

Teacher reprimands were defined generally (e.g., “Stop!”; “Don’t do that!”; “What are you supposed to be doing?”), including directed reprimands (e.g., “This classroom is too loud!”; “Group A, you should be sitting.”; “Kit, you just lost a point.”; “Tai, mind your business.”) and behavior specific reprimands (e.g., “Johnny, you are being too loud!”; “Jeremiah, why are you standing up here?”; “Kit, you just lost a point.”). Specific non-directed reprimands were also included (“This classroom is too loud and is about to lose a point.”). Reprimands also included instances in which points were deducted without a remark by the teacher for inappropriate behavior within the context of the GBG. Non-examples included redirections and demands (e.g., “Jay, sit down.”), and reprimand points or reprimands provided by someone other than the

teacher. All three types of teacher reprimands collapsed into a single frequency count of reprimands.

Student On-Task Behavior

The definition for on-task behavior was defined as those times when the student was attending to the assigned work (e.g., writing, reading aloud, raising a hand, talking to peer about assigned material) or passively attending to assigned work (e.g., looking at the teacher as she speaks to the class, reading assigned material silently). Non-examples included walking around the classroom without permission, calling out, aimlessly looking around the classroom, and silently reading unassigned material; this definition was adapted from Riley-Tilman and colleagues (2009). This definition was printed on the Direct Behavior Rating (DBR) rating sheet provided to the teachers.

Procedure

Baseline

Baseline sessions were conducted during a 20-minute observation period that the teacher identified as the most disruptive. Teachers were instructed to conduct classroom management in the typical manner, including any reinforcement strategies they usually implement. At least three sessions of praise and reprimand data were collected without the GBG in place to record the natural rate of teacher behavior. This determined the variable-interval schedules for the intervention phases.

Good Behavior Game Teacher Training

After baseline, the teachers were trained to implement the GBG based on the teacher script (Appendix E) and the treatment integrity checklist (Appendix I), which allowed for consistency of teacher implementation of the GBG across sessions. The

components of the GBG were trained using behavioral skills training, which included didactics, modeling, role-play, and performance feedback. Teachers moved onto the GBG baseline phase upon 100% mastery of all GBG steps.

Assessment of procedural integrity occurred during all initial training sessions with a checklist (Appendix J; Lambert, Tingstrom, Sterling, Dufrene, & Lynne, 2015; Ford, 2015). Each training session resulted in 100% procedural integrity scores.

Good Behavior Game Baseline

Once the baseline was completed, there were at least three sessions of praise and reprimand data collected with the GBG in place. At this point, the schedules of reinforcement for the teacher behavior were not manipulated by the primary investigator. This phase's purpose was to verify if rates of praise and reprimands changed due to the implementation of the GBG. Secondly, the rates of praise and reprimands for this phase determined the VI schedules for the intervention phases.

Teacher Training

The primary researcher conducted a brief training with each teacher between the GBG baseline and intervention phases. Teachers were told that they had an opportunity to earn a reward based on the number of times they award and remove points from their students within the context of the GBG. Specifically, they were told that their behaviors would be rewarded for making the GBG as "active" as possible by frequently providing or removing points from teams when classroom rules are followed or broken, respectively.

Intervention

For the first intervention phase (aka praise phase), teacher praise and reprimands were placed on separate concurrent variable interval-variable interval (VI-VI) reinforcement schedules. The proportion of reinforcement available for both schedules was derived from the rates of praise and reprimands observed in the implementation baseline sessions. If a teacher praised students twice as much as they reprimanded students the VI schedule for praise was twice as dense as the VI schedule for reprimands (e.g., VI 30s, VI 60s). Due to the way the GBG was established, teachers had the discretion to add or remove points for student behavior at any time. Each time the teacher satisfied one of the VI schedules, they earned a point, displayed as a tally mark on a small dry erase board at the back of the classroom. The primary investigator maintained this dry erase board, tallying the total number of praise statements and reprimands delivered during each session. At the end of each session, an observer showed the teacher the number of points they earned for that session, which allowed them to access rewards through a token economy. For example, if a teacher indicated that she preferred black Expo markers, after earning 38 points, she could obtain an EXPO marker.

Reprimand Phase

In the next phase (aka the reprimand phase), everything remained the same as above, with one exception. In the reprimand phase, the proportion of reinforcement available on both VI schedules was flipped. For example, if the VI schedule for praise statements was originally set on a 30s interval and the VI schedule for reprimands was originally set on a 60s interval, the two values (i.e., 30s and 60s) were swapped, making the VI schedule for reprimands a VI 30s schedule and the VI schedule for praise

statements a VI 60s schedule. Classroom C was not included in this or the following phases due to time constraints in the spring semester.

Balanced Phase

Lastly, a phase in which the proportion of reinforcement available on both VI schedules was equal. Both the VI schedule for praise and the VI schedule for reprimands were set at an interval that equaled the average interval length between the two during the GBG baseline phase. For Classroom A, this was VI 37; for Classroom B, this was VI 29.

Interobserver Agreement

IOA data were collected by graduate students who completed training with the primary investigator on data collection for this study. All data collectors also completed Collaborative Institutional Training Initiative (CITI) training on Human Subjects Research prior to data collection. IOA was calculated for at least 25% of all sessions using the mean count per interval IOA. Praise statement IOA was calculated by dividing up the data into ten second intervals, recording the number of occurrences of behavior within each interval, calculating the agreement between each observer within each interval, and multiplying by 100 (i.e., (interval 1 IOA + interval 2 IOA… + interval N IOA/ n intervals) * 100). The same was done for reprimands. The minimum acceptable IOA was 85%. If a datum fell below this criterion, feedback was provided to the data collector. If the datum fell below the criterion a second time, the data collector was re-trained. However, no IOA datum fell below 85%, so re-training was unnecessary.

Overall, Classroom A had a percentage of 32% of sessions with IOA. This included 33.33% of baseline sessions. Forty percent of GBG baseline sessions had IOA data collection. IOA sessions per phase included 16.67%, 50%, and 25% for intervention

phases, respectively. Across phases, IOA for teacher praise statements ranged from 85.91-96.82, averaging 93.94%. For teacher reprimands, IOA ranged from 85.53-99.09, averaging 95.41%.

Classroom B's overall IOA percentage of sessions was 29.03% of sessions. This included 33.33% of baseline sessions and 33.33% of GBG baseline sessions. IOA sessions included 20% of Praise VI sessions and 66.67% of Equal VI Phase sessions for intervention phases. Zero percent of sessions in the reprimand phase had IOA due to Finals Week for graduate students, which reduced the availability of data collectors. Across phases, IOA for teacher praise statements ranged from 89.55-99.09, averaging 94.63%. For teacher reprimands, IOA ranged from 85.23-100, averaging 93.93%.

The total percentage of IOA sessions for Classroom C was 36.36%. IOA was collected for both baseline and GBG baseline phases for 33.33% of observations. For the praise phase, 50% of sessions included IOA, which was one session. This session's IOA for teacher praise statements was 88.89%, and teacher reprimand IOA was 95.91%.

Treatment Integrity

Treatment integrity was assessed during the observation periods with a treatment integrity checklist completed by observers. The checklist assessed if the teacher implemented the GBG correctly for 100% of sessions within the GBG baseline and all intervention phases across all three classrooms. The treatment integrity checklist consisted of 9 "yes," "no," or "N/A" statements. One example of an integrity step was, "Teacher allows winning team(s) access to the reward," which is also a step that could have resulted in a "N/A" if none of the teams won the game. Treatment integrity was calculated by dividing the number of steps the teacher completed correctly and dividing it

by the total number of steps on the checklist. This number was multiplied by 100 to obtain a percentage for the accuracy of steps for each teacher per session. IOA for treatment integrity occurred during at least 25% of behavioral observations (across classrooms and phases). If treatment integrity decreased to fewer than 80%, the teacher was re-trained. Following each observation, the teacher was provided performance feedback regarding treatment integrity (Noell et al., 2005). For Classroom A, treatment integrity averaged 71.12% across phases, ranging from 66.7-100%. In session ten, the teacher's integrity was 66.7%; therefore, the teacher was retrained on all steps of the GBG. For Classroom B, treatment integrity averaged 97.33%, ranging from 88.89-100%. For Classroom C, treatment integrity averaged 58.18%, ranging from 11.11-100%. In session 8, the teacher's integrity was 11.11%, so she was retrained on the entirety of the GBG. This was the only session that required retraining. Anecdotally, after the observation, the teacher stated it was a rough day, and she did not have her script. Her median integrity percentage was 88.89%. IOA for all integrity sessions for all teachers was 100%.

Data Analysis

Generalized Matching Equation ($\log(B1/B2) = a \log(R1/R2) + \log b$) was used to predict if teachers allocated their responses according to the reinforcement schedule or if overmatching or undermatching occurred. Each type of behavior response was represented as B1 and B2 and were recorded as frequencies. R1 and R2 were the relative rates of reinforcement the teacher received from each reprimand or praise statement. *a* reflected the slope and sensitivity to relative reinforcement rates (if any). *a* measures the slope's best fit line; if *a* is less than one, this suggested that the teacher maximized the

available rate of reinforcement (Reed, 2009). *b* reflected the bias (i.e., preference) of an alternative over the other (Borrero et al., 2010) that was not attributable to reward points obtained. A bias greater than zero suggested a bias for the first alternative (B1/teacher praise), whereas a bias less than zero suggested a bias for the second alternative behavior (B2/teacher reprimand) (Reed, 2009). If slope *b* was greater than zero, there was a positive bias for the response in the numerator of the GME equation. If slope *b* was less than zero (a negative bias), then there was a bias for the response in the denominator. If the number obtained for bias was equal to or near zero, it was ascertained that no bias was present.

The GME was chosen over other matching equation equations because it accounts for variations of strict matching by integrating sensitivity and bias into the equation. As one may expect in a natural environment, there are countless competing contingencies in play. A teacher may be biased or sensitive in choosing reprimands versus praise that is not accounted for by reinforcement. Hence, the researchers accounted for that with *a* and *b* in the GME equation.

The GME utilizes linear regression analysis, which includes calculating the best fit line, which allows for the calculation of R^2. R^2 is the variance accounted for and goodness of fit for by the GME in relation to the data. R^2 results range from 0.0-1.0. The closer a value to 1.0, the better match of the data to the GME. The goodness of fit values can also be interpreted as a percentage of the variance explained by the GME. Percentage values closer to 100 indicate a closer match to the GME (Reed, 2011). R^2 was computed instead of other single case effect size calculations due to linear regression to calculate the GME.

CHAPTER III - RESULTS

Teacher Behavior

Baselines

Classroom A Baselines

Rates of both praise and reprimand statements were calculated by dividing the number of times the behavior occurred by the 20-minute observation period; thus, resulting in the rate per minute of the behaviors (reported in Figure 1). Classroom A's praise statement rates averaged 0.38 praises per minute during baseline and ranged from 0.10-0.50 per minute (i.e., the frequency range of 2-10 total praises). The reprimand average rate was higher than praise with 0.87 per minute (range: 0.5-1.3 per minute). In contrast, upon implementing the GBG, praise statements increased to an average of 1.62 per minute (range: 1.4-2.2 per minute), while the reprimands' average was 0.5 per minute (range: 0.3-0.9 per minute). Of note, there were no overlapping data points in Classroom A's teacher behavior. Based on the results of the GBG baseline, on average, praise occurred every 37.04 seconds, with an SD of 6.45. A reprimand occurred on average every 120 seconds, with an SD of 54.18.

Classroom B Baselines

In Classroom B's baseline, overall, there was a higher rate of praises than reprimands, with a decreasing trend for both reprimands and praises in session six. When asked if there was anything different about that day, the teacher described it as a "hard day." Praise statement rates occurred an average of 1.5 per minute (range: 0.4-2.1), while reprimand rates average was 1.15 per minute (range: 0.6-1.7 per minute). In the second baseline phase (GBG), there were zero overlapping data points. Praise rates ranged from

0.9-3 per minute, averaging 2.05 per minute; thus, praise statements were typically higher than in the baseline. Reprimand rates in this phase ranged from 0-0.8 per minute, with an average of 0.3 per minute. Based on the results of the GBG baseline, praise occurred every 29.27 seconds, with an SD of 12.53. The average reprimand occurred every 189.47 seconds, with an SD of 153.17.

Classroom C Baselines

The baseline in Classroom C included an average praise rate of 0.4 per minute (range: 0.1-0.6 per minute); C's average reprimand rate was 0.95 per minute (range: 0.5-1.9 per minute). The higher reprimand rates for inappropriate behavior are typical of teacher behavior (Beaman & Wheldall, 2000). In the GBG baseline phase, rates of teacher behavior were variable. Praise rates averaged 1.6 per minute (range: 0.9-2 per minute), while reprimands rates averaged 1.2 per minute (range:0.9-1.6 per minute). In other words, praise occurred every 38.3 seconds on average (SD=20.28), and reprimands happened every 51.43 seconds on average (SD=12.59).

Praise Phases

Classroom A Praise Phase

Within the first intervention phase, the VI schedule for Classroom A was VI 37 seconds for praise and VI 120 seconds for reprimands. Graphs on rates (Figure 1) showed that teacher behavior continued to have zero over-lapping data points. A range of 1-1.9 praises per minute occurred, with an average of 1.4 per minute. However, 0.4 average reprimands occurred per minute, with a range of 0.1-0.6 per minute. Sessions 12, 13, and 17 included reprimands occurring at the rate of 0.5 per minute. Due to the VI schedule, the primary investigator hypothesized that praise reinforcement would happen with an

average rate of 1.62 per minute. In contrast, reprimand reinforcement would occur at 0.5 per minute on average. The range of praise reinforcement was 0.5-0.8, averaging 0.7 per minute. The range of reprimand reinforcement was 0.1-0.3, averaging 0.21 per minute. None of the sessions obtained the average rates for perfect matching, although praise rates and reinforcement were higher than reprimand rates and reinforcement.

The first intervention phase's GME data is found in Table 1 for Classroom A. In six sessions, the reinforcement sensitivity was 1.505, so overmatching occurred. This suggested that the rate of behaviors was more than required to obtain reinforcement. Since the reinforcement rate was higher for praise statements than reprimands, one may hypothesize a bias occurred for praises. This did not happen as predicted, as bias equaled -0.1310, which means a bias for reprimands. Of note, higher praise reinforcement rates are not necessarily indicative of a higher rate in matching. However, the best fit line indicated that 73.10% of the variance might be accounted for by the GME equation, given the known relative reinforcement rates.

Classroom B Praise Phase

Within the first intervention phase, due to the rate of 1.5 per minute average in GBG baseline, the VI schedule for Classroom B was VI 29 seconds for praise (rate of 2.05 per minute) and VI 189 seconds for reprimands (rate of 0.3 per minute). Praise ranged from 1.6-2.3 per minute, with an average of 1.94 per minute. Session 22 had 2.3 praises per minute, which was the closest session in the praise phase to reach 2.05 per minute. Praise reinforcement rates ranged from 0.7-0.85, averaging 0.74 per minute. Reprimands occurred with a range of 0.1-0.5 per minute, averaging 0.3 per minute. Session 19 was the only session in which reprimands occurred at a rate of 0.3 per minute.

Reprimand reinforcement rates ranged from 0.05-0.2 per minute, averaging 0.12 per minute (Figure 1). Despite the VI schedules in place, Classroom B continued to have higher rates of praise statements.

In Classroom B's first intervention phase, the reinforcement sensitivity was 0.5602, which indicates undermatching. This means that when there was an increase in reinforcement, the increase in behavior was less than predicted (i.e., 0.5602). Therefore, the teacher's behavior did not change because of the reinforcement rates. With a y-intercept or bias of 0.7031, Classroom B was biased towards providing praise statements rather than reprimands, which was hypothesized since relative reinforcement rates were more available for praise statements in this phase. The value of the best fit line equaled 0.09226, which means that only 9.2% of the behavior pattern could be explained by the GME, given the relative rates of reinforcement in this phase. Therefore, the GME did not account for much of the behavior pattern, and other variables were at play outside of the relative rates of behavior and reinforcement.

Classroom C Praise Phase

The requirement for Classroom C's teacher to receive a reward(s) was meeting the schedules of VI 38 seconds for praises (rate of 0.6 per minute) and VI 51 seconds for reprimands (0.9 per minute). In the two sessions completed, 2.10 and 2.0 praises occurred per minute in sessions 10 and 11, respectively. Praise reinforcement rates were 0.8 per minute for both sessions. Although Classroom C only had two sessions, her behavior had the most significant gap between rates of reprimands and praise.

While praise statements increased in the praise phase, the statements did not match the available reinforcement rates. In contrast, 0.4 and 0.5 reprimands occurred per

minute, whereas reprimand reinforcement rates were 0.35 and 0.45 per minute (Figure 1). While a decrease in rates of reprimands from baseline is typically revered in the research literature, in this case, the teacher's behavior did not match the rates of reinforcement available.

Classroom C's sensitivity to reinforcement equaled 2.952, which means that the teacher's behavior rates were much more than required to receive reinforcement. In this case, when the relative rate of reinforcement was one, the teacher responded more than three times as much as necessary on average than required. Although praise rates and praise reinforcement were higher in Classroom C, bias was -0.3398, indicating a bias for reprimands. Variance was 0.9502, so about 95% of the variance within the data may be explained by the GME. Due to end-of-the-semester time constraints, the teacher chose not to proceed with the study after session 11; therefore, no data were collected for reversal or balanced phases for Classroom C.

Reprimand Phases

Classroom A Reprimand Phase

Within the second intervention phase, the VI schedule for Classroom A was VI 120 seconds for praise, and VI 37 seconds for reprimands. Due to the VI schedule, it was hypothesized that praise reinforcement would occur with an average rate of 0.5 per minute, while reprimand reinforcement would occur at 1.62 per minute on average. However, higher rates of praise statements continued to occur. Praise rates ranged from 1.3-1.6 per minute, averaging 1.43 per minute. Praise reinforcement ranged from 0.3-0.4, with an average of 0.33 per minute. In contrast, reprimand rates ranged from 0.1-0.4 per

minute, averaging 0.23 per minute, whereas reprimand reinforcement rates ranged from 0.1-0.3, averaging 0.2 per minute (see Figure 1 for graphs).

In Classroom A's reprimand phase, the GME data described a 0.8992 sensitivity to reinforcement (Table 1), revealing undermatching occurred. Thus, fewer responses were emitted than required or predicted by the VI schedule. This is also indicative of responses allocated to praises, which was further supported since bias was 0.6402. Therefore, the teacher allocated more responses to praise statements rather than reprimands. R squared was 0.9776, so the GME accounted for 97.76% of the variance.

Classroom B Reprimand Phase

The VI schedule for Classroom B was VI 189 seconds for praise statements (rate of 0.3 per minute) and VI 29 seconds for reprimands (rate of 2.05 per minute). However, the teacher continued to allocate more responses to praise statements. Praise rates ranged from 1-1.6 per minute, with an average of 1.32 per minute. Praise reinforcement rates ranged from 0.1-0.2 per minute, averaging 0.15 per minute. Reprimand rates ranged from 0.2-0.3 per minute, with an average of 0.28 per minute. Four out of the five sessions had rates of 0.3 reprimands per minute. However, reprimand reinforcement rates ranged from 0.15-0.25 per minute, averaging 0.19 per minute. Even though the teacher had the expected rate of reprimands per minute, she did not respond to the VI schedule.

In the GME analysis for Classroom B's reprimand phase, reinforcement sensitivity was 0.4891, indicating that undermatching occurred. Therefore, fewer responses were made than the available reinforcement. Bias equaled 0.7872; hence, the teacher allocated more responses to praise statements, despite the availability of a thicker schedule of reinforcement for reprimands. Of note, R squared was 0.5042, so the GME

only explained about 50% of the variance within the data. The other 50% of the variance was explained with variables outside of the GME, which is plausible due to the bias of praise statements despite the VI schedule.

Balanced Phases

Classroom A Balanced Phase

The schedule was VI 37 seconds for both praise and reprimand statements. Overall, there was a downward trend for praise and a slight increase for reprimands. Anecdotally, the researchers observed that Teacher A would look towards the observers and researcher's board more often in this phase. Praise statements ranged from 0.95-1.9 per minute, with an average of 1.39 per minute. Praise reinforcement ranged from 0.6-0.8 per minute, averaging 0.7 per minute. In contrast, reprimands ranged from 0.3-0.6 per minute, with an average of 0.34 per minute. Reprimand reinforcement ranged from 0.1-0.3, averaging 0.23 per minute.

In Classroom A's balanced phase, the GME described a 1.225 sensitivity to reinforcement (Table 1), indicating overmatching occurred. Moreover, bias was -0.01110, meaning there was essentially no bias in responding. This is the first, and only time no bias occurred within the study. Thus, when reinforcement rates for praise and reprimand statements were equal, the teacher's preference for behavior was likely based on reinforcement rates alone. This is further confirmed by 95.72% of the variance being accounted for by the GME.

Classroom B Balanced Phase

The VI schedule for both praise and reprimand statements was VI 29 seconds (i.e., 2.05 per minute). Praise statements ranged from 1.1-1.4 per minute, with an average

of 1.27 per minute. Praise reinforcement ranged from 0.65-0.75 per minute, averaging 0.7 per minute. In contrast, reprimands ranged from 0.6-0.8 per minute, with an average of 0.67 per minute, closer to the reprimand rates in the baseline. There was also less of a gap displayed between the praise and reprimand data sets within this phase compared to other intervention phases. Lastly, reprimand reinforcement ranged from 0.3-0.55, averaging 0.42 per minute.

In the GME analysis for Classroom B's balanced phase, reinforcement sensitivity was 0.7134, revealing that undermatching occurred. Bias data suggested a bias for praise (b = 0.1110). R squared was 0.8718, indicating that 87.18% of the variance within the data may be explained by the GME rather than extraneous variables.

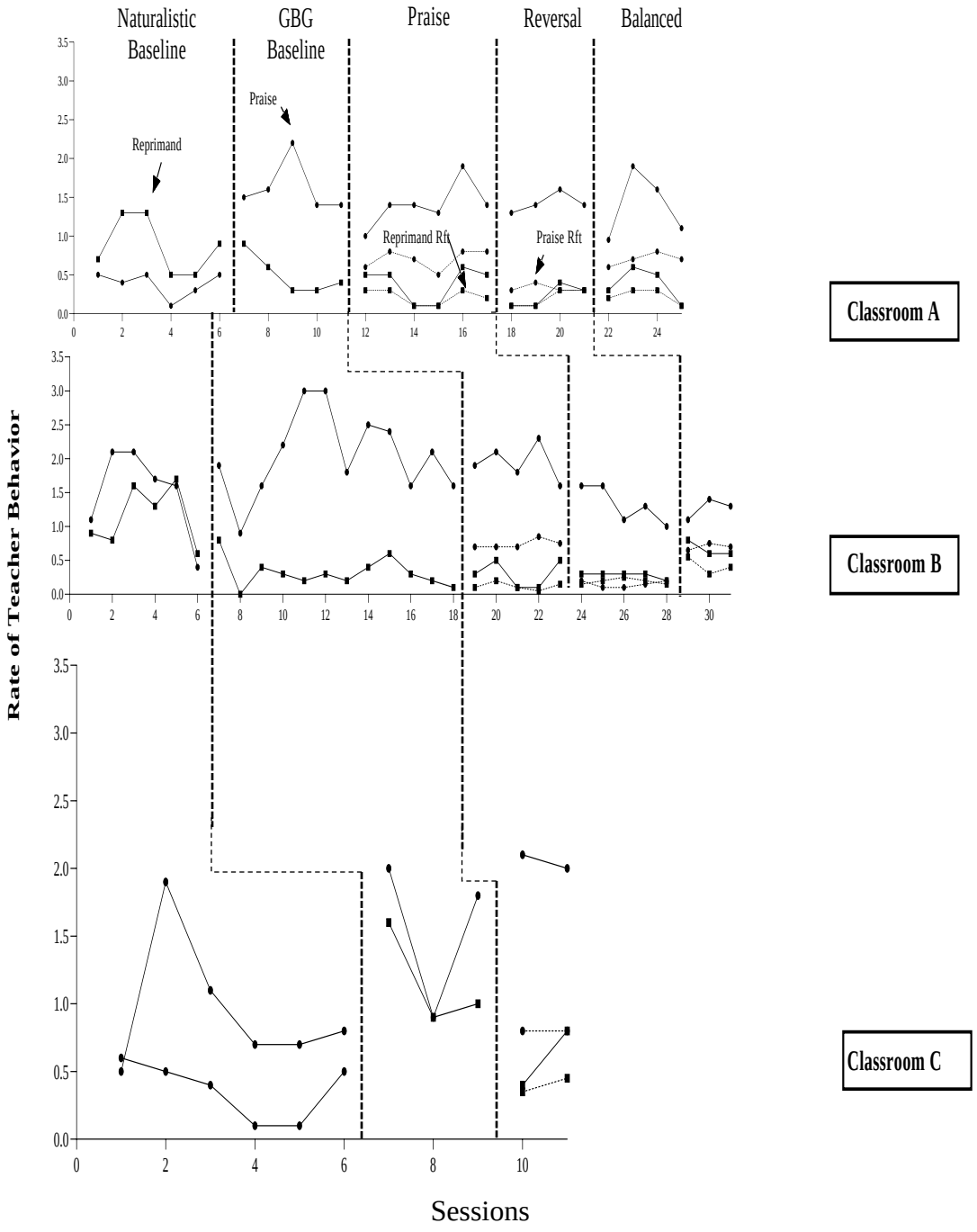
Table 1

Generalized Matching Equation (GME) Results

Classroom	Sensitivity to Reinforcement	Bias	Variance Explained
	Praise Phase		
A	Overmatching	Reprimands	73.10%
B	Undermatching	Praises	9.2%
C	Overmatching	Reprimands	95.02%
	Reprimand Phase		
A	Undermatching	Praises	97.76%
B	Undermatching	Praises	50.42%
C	N/A	N/A	N/A
	Balanced Phase		
A	Overmatching	No Bias	95.72%
B	Undermatching	Praises	97.18%
C	N/A	N/A	N/A

Note. The above table represents the comprehensive results of the Generalized Matching Equation data analysis across classrooms. "N/A" or "Not applicable" is noted in Classroom C, meaning Classroom C did not participate in the Reprimand or Balanced Phases of the current study.

Figure 1.



DBR Student Behavior

For Classroom A, during baseline, on-task behavior averaged 50%, ranging from 30-70%. In the GBG baseline, ratings averaged 78%, with each session rated as either

70% or 80% on-task. Within the intervention phases, data were as follows: for the praise phase, ratings averaged 85%, ranging from 80-90%; for the reprimand phase, ratings averaged 72.5%, ranging from 60-90%; for the Equal VI Phase, ratings averaged 77.5%, ranging from 70-80%.

For Classroom B, during baseline, the on-task behavior averaged 56.67% ranging from 40-70%. During the GBG baseline, ratings increased and averaged 89.17%, ranging from 80-100%. Within the intervention phases, on-task behavior ratings were as follows: for the praise phase, ratings averaged 90%, with an 80-100% range. For the reprimand phase, ratings averaged 92%, ranging from 90-100%. In the last phase, ratings averaged 86.67%, ranging from 80-90%.

For Classroom C, baseline ratings averaged 40% on-task for class-wide behavior, ranging from 20-50%. In the GBG baseline, on-task behavior ratings averaged 70%, ranging from 50-90%. In its single intervention phase, ratings averaged 85%, ranging from 80-90% (see Table 2 for a summary of DBR results).

Table 2

Across Phases: DBR Results

Classroom	Baseline	GBG Baseline	Praise Phase	Reprimand Phase	Balanced Phase
A	50%	78%	85%	72.5%	77.5%
B	56.67%	89.17%	90%	92%	86.7%
C	40%	70%	85%	N/A	N/A

Note. The percentages reported above are averaged across each phase per classroom (i.e., for Classroom A in the Baseline, students were reported as on-task 50% of observations, on average. N/A is noted in Classroom C due to the classroom completing three of five phases.

CHAPTER IV - DISCUSSION

The purpose of this study was to extend the research on the generalized matching equation into an applied setting to determine if teachers would allocate their responses to concurrent reinforcement rates. This was evaluated by analyzing the rates of teacher behavior compared to the reinforcement available, along with assessing bias and sensitivity to reinforcement. It was hypothesized that teachers would allocate their responding according to either the reinforcement schedule and/or student behavior. Limitations and future directions based on this study are also discussed.

Research Questions

Questions 1 and 2

The first research question addressed whether the proportion of teachers' praise statements to reprimands would alter when a contrived contingency was implemented that conforms to the generalized matching equation, or were biases present? This is explained in part by the rates of responses and rates of reinforcement and further explained within the GME analysis. The second research question investigated the extent to which the proportion of teachers' praise statements to reprimands match, undermatch, or overmatch the contrived contingency when the proportion of reinforcement is modified. This question is answered within the GME analysis.

Classrooms A and B both had praise and praise reinforcement rates higher than reprimand rates and reinforcement rates in the praise phase, with no overlap between praise and reinforcement data. Overall, Classroom C exhibited higher praise and praise reinforcement rates, with some overlap. However, since overmatching occurred for Classroom A, undermatching in Classroom B, and overmatching in Classroom C,

teachers' allocation of responding did not match the GME. In addition, each teacher exhibited biases in their responding. There were biases for reprimands for Classrooms A and C, whereas Classroom B had a bias for praise statements. Lastly, the variance explained for Classroom A was 73.10%, whereas the GME explained only 9.2% of the variance for Classroom B. Classroom B's variance in the first intervention phase was the lowest across the study. In contrast, Classroom C's R^2 value was very high, with 95.02% of the variance explained by the GME. In most studies on human behavior, any variance higher than 50% is unlikely due to the complications of human behavior; however, due to the factors for bias and sensitivity to reinforcement, this may be more likely with behaviors and reinforcement rates under the influence of the GME calculation.

Classrooms A and B continued to have higher rates of praise data in the reprimand phase than reprimands, suggesting a carryover effect from the first intervention phase. Furthermore, in Classroom A, average praise rates were almost identical for both phases (1.4 and 1.3, respectively). Average reinforcement for praise decreased from 0.7 per minute in the first intervention phase to 0.3 per minute as predicted-perhaps in part due to the decrease in the availability of reinforcement for praise statements. The average reprimand reinforcement also remained consistent from the previous intervention phase, which was 0.21 per minute in the first intervention phase and 0.23 per minute in the reprimand phase. This indicated that although praise reinforcement rates decreased, reprimand reinforcement rates did not decrease as expected.

In addition, Classroom B had the steadiest data trend (downward) in the reprimand phase, which is expected due to the variable interval schedule. However, only

session 19 included the expected rate of 0.3 reprimands per minute, with an average of 0.12 reprimands reinforced by observers per minute according to the VI schedule. None of the sessions included an expected praise rate of 2.05 per minute.

Classroom A and B's GME analyses determined that undermatching occurred in the reprimand phase, with biases for allocating responding to praise statements. The biases for praise are the opposite of GME predictions due to the availability of reinforcement being higher for reprimands. This further added to the likelihood of carryover effects from the previous phase. In Classroom A, 97.76% variance explained, whereas, in Classroom B, 50.42% of variance explained.

In the final intervention phase, the availability of reinforcement was equal for praise and reprimand statements (i.e., balanced). Interestingly, in Classroom A, the average praise reinforcement rate was the same in the praise intervention phase and the balanced phase (i.e., 0.7 per minute). Classroom A's GME analysis indicated overmatching, although bias was nearly zero (which only happened in this study phase). A zero bias matches the hypothesis for the balanced phase of the study. Plus, since the GME explained 95.72% of variance, conclusions may be confidently drawn from those data. However, Classroom B's data did not show equal allocation of responding to praise and reprimand statements. The average praise rate per session remained higher than the average reprimand rate per session. Overall, undermatching occurred with a bias for delivering praises; furthermore, 87.18% of the variance was explained by the GME.

Overall, these results match the current GME literature within the classroom (Davison & McCarthy, 2016; Billington & DiTomasso, 2003). Previous studies in the behavior allocation literature have questioned potential reasons for bias and sensitivity in

behavior, as were seen in the current study. Potential reasons may include the immediacy of reinforcement, alternatives for high-quality reinforcement (Billington & DiTomasso, 2003), and discriminability between reinforcers (Davison & McCarthy, 2016). Regarding the current study, teachers were delayed in obtaining rewards. They chose their rewards at the end of the twenty-minute observation. If they did not have enough points, they had to wait until the end of another observation to collect a reward. This caused even more of a delay in obtaining items. Additionally, the quality of rewards may have been an issue. Principals only approved school supply items, and the rewards may have had less value. Also, indiscriminability between reinforcers was likely. Tally points were undifferentiated regarding what was written on the whiteboards (i.e., color, side of the board). Teachers may need more salient prompts, due to multiple concurrent schedules in the classroom.

Lastly, in comparing the DBR data for class-wide on-task student behavior across classrooms, there were increasing trends for all classrooms (A-C) from baseline to GBG baseline and the first intervention phase. This included some carryover from GBG baseline to the first intervention phase for only Classrooms A and B, which is unsurprising given the carryover of higher rates of praise statements in both classrooms. Classroom C's DBR data continued to increase without carryover into the first intervention phase.

In the reprimand phase, given the literature on correlations between inappropriate teacher behavior (e.g., reprimands) and off- task student behavior (Kodak, Miltenberger, & Romaniuk, 2003), one may hypothesize a decrease in the DBR percentage. This held true for Classroom A, with a range of 60-90% on-task ratings, averaging 72.5%,

compared to a previous average of 85% class-wide on-task behaviors in the previous phase. However, for Classroom B, the average on-task ratings continued to rise to an average of 92% class-wide. Since the teacher in Classroom B's responding was biased towards praise statements, this makes sense. Due to that bias, she may have encountered a resurgence of inappropriate behavior, although it should be noted that after the GBG baseline, her ratings never fell below 80%. A rating of 80% did occur during the reprimand phase, so it is possible that the teacher perceived 80% on-task as low.

In the balanced phase, the ratings for class-wide on-task behavior increased from an average of 72.5% in the previous phase to an average of 77.5%. This paralleled with Classroom A's lack of bias in praise versus reprimand statements since that average was between the ratings for the praise and reprimand phases. However, Classroom B's lowest ratings for intervention occurred in the balanced phase, averaging 86.67%, which falls below the GBG baseline ratings. The percentage is higher than the baseline, which averaged 56.67%.

Limitations

Several limitations should be considered when evaluating this study. There were some indications of rule-governed contingencies at play for the delivery of praise. The sufficient change in appropriate student behavior from the baseline to the GBG baseline may have been enough to compete and render the VI schedules irrelevant. It is possible that when Classroom B increased rates of reprimands even slightly, the DBR data did not reflect the higher on-task student behavior; therefore, increases in on-task behavior were not salient enough for the teacher to differentiate changes in student behavior or reinforcement rate provided by the data collectors. However, this may be mollified with

Classroom A and C's biases for reprimands in the first intervention phase and Classroom A's bias for neither praise nor reprimands in the balanced phase. In addition, due to the delay in reinforcement with school supplies, it was likely that the items were not functioning as effective reinforcers. A more powerful reinforcer was competing, such as student behavior.

Future Directions

It appeared for the current study that the rewards were not potent enough relative to the direct acting contingencies in place related to immediate consequences of praise and reprimands of student behavior on teachers. Hypotheses pertaining to the allocation of student behavior contributions to allocations of teacher behavior should be addressed in future studies, or perhaps the extent of rule governed behavior should be tackled. Although several extraneous variables are possible due to concurrent schedules and rule governed behaviors, future studies may likely test these hypotheses. One such way would be to gain parent consents for videos of student behavior to determine the extent of the direct acting contingencies on teacher behavior and student behavior.

Future studies may employ various ways to control for carryover of praise statements. One such option would be to employ a different VI schedule after the GBG baseline. Instead of allotting a thicker schedule of reinforcement for praise statements, allot more reinforcement for reprimands first. Another option would be only to involve one version of the GBG instead of the combined version in which teachers may add or deduct points based on student behavior.

Additionally, regarding rewards utilized, delays in provision may be addressed by lying out items on a table as they are earned. A formal preference assessment may assist

in the selection of such items instead of a reward menu. This may increase the saliency of the delivery of reinforcement, although a balance must simultaneously decrease the likelihood of rule-governed contingencies. Future studies may also employ exclusionary data when recruiting teacher participants. Classroom B likely did not require the GBG given that the teacher perceived class-wide on-task behavior high in the GBG baseline.

Conclusion

Findings for Classroom B align with those of Elswick and Casey (2011) and Lannie and McCurdy (2007) which found that when given a choice between adding and removing points for the GBG, teachers may allocate responding to the addition of GBG points for appropriate behavior; in addition, this study adds that when given a choice, teachers may allocate their responding to provision of praise rather than reprimands, as the teacher in Classroom B throughout the study. The data for this study confirm that teachers may not be sensitive to the contingencies at play (Kodak, Miltenberger, & Romaniuk, 2003), but further research is necessary into the effects of student behavior on teacher behavior. However, the findings for Classroom A's balanced phase allude to some teachers' capabilities of allotting behaviors to a reinforcement schedule within the concurrent schedules of the classroom setting. More research on the GME in the classroom environment is warranted to confirm these findings.

APPENDIX A – Teacher Demographic Form

Name: _____			**Current Grade(s) Taught: _____**	
Please indicate (circle) the following:				
Identify your gender:	**Identify your ethnicity:**	**Identify your age range:**	**Select your education level:**	**Identify your level of teaching experience:**
Male	African-American	1945-1960	Bachelor's Degree	10 or more years
Female	Caucasian	1961-1979	Master's Degree	5 or more years
Other	Hispanic	1980-1994	Doctorate Degree	2 or more years
	Other	1995-present		Less than 2 years

APPENDIX B – IRB Approval Letter

THE UNIVERSITY OF
SOUTHERN MISSISSIPPI

INSTITUTIONAL REVIEW BOARD
118 College Drive #5147 | Hattiesburg, MS 39406-0001
Phone: 601.266.5997 | Fax: 601.266.4377 | www.usm.edu/research/institutional.review.board

NOTICE OF COMMITTEE ACTION

The project has been reviewed by The University of Southern Mississippi Institutional Review Board in accordance with Federal Drug Administration regulations (21 CFR 26, 111), Department of Health and Human Services (45 CFR Part 46), and university guidelines to ensure adherence to the following criteria:

- The risks to subjects are minimized.
- The risks to subjects are reasonable in relation to the anticipated benefits.
- The selection of subjects is equitable.
- Informed consent is adequate and appropriately documented.
- Where appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of the subjects.
- Where appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of all data.
- Appropriate additional safeguards have been included to protect vulnerable subjects.
- Any unanticipated, serious, or continuing problems encountered regarding risks to subjects must be reported immediately, but not later than 10 days following the event. This should be reported to the IRB Office via the "Adverse Effect Report Form".
- If approved, the maximum period of approval is limited to twelve months.
  Projects that exceed this period must submit an application for renewal or continuation.

PROTOCOL NUMBER: 18020604
PROJECT TITLE: Matching Equation: Teacher Rates of Praise and Reprimands
PROJECT TYPE: Master's Thesis
RESEARCHER(S): Meleah Ackley
COLLEGE/DIVISION: College of Education and Psychology
DEPARTMENT: Psychology
FUNDING AGENCY/SPONSOR: N/A
IRB COMMITTEE ACTION: Expedited Review Approval
PERIOD OF APPROVAL: 06/05/2018 to 06/04/2019

Edward L. Goshorn, Ph.D.
Institutional Review Board

APPENDIX C – Teacher Consent Form

Dear Teacher,

Hello, my name is Meleah Ackley, and I am a graduate student at the University of Southern Mississippi in the School Psychology Doctoral Program. I am currently conducting my thesis, which will assess the effectiveness of a classroom behavioral intervention. This study is being conducted under the supervision of Dr. Evan Dart.

Please consider the following when deciding if you will participate in this study:

Purpose of the Study

The purpose of this study is to assess the effectiveness of a class-wide intervention known as the GBG in relation to classroom rules. The GBG utilizes classroom management to increase rule following and decrease rule violations.

Procedure

If you agree to participate in this study, you will be asked to perform various tasks associated with classroom management. Prior to implementation of the intervention, it is required that you complete consultation and training sessions with me. After the consultation session, a series of screening observations will determine if your classroom is appropriate for this study. During that time, you will be asked to follow your normal classroom management techniques for problem behavior.

If your classroom qualifies, you will be asked to complete a training session with me that explains the steps of the GBG and practice the intervention. After you are able to complete the steps with 100% accuracy and describe each step, we may begin the intervention.

During the intervention, you will be asked to add points based on rule following and remove points based on rule violations. You may use the white board to tally marks for each team. The goal of the game is for students to earn greater than or equal to a preset criterion. Contingent on meeting that criterion, the team gains access to a preferred reward. On each day of the intervention, you will be asked to state the rules (a script will be provided for you to read from), award and remove marks as appropriate, state the criterion, and hand out rewards (which will be provided) to the winning team(s).

Through the course of the study, there will be classroom observations a few times each week. Observations will be conducted by trained graduate students from the USM School Psychology Program. If all students' guardians consent to video recordings of student behavior, you also consent to the video recording. If all guardians do not consent, videos will not be recorded. As necessary, you will be provided feedback on the implementation of the game throughout the study.

Benefits

By agreeing to this study, there may be several benefits for you and your students. You will be trained in implementing an empirically validated classroom management technique that can be used with your current and future students. Additionally, there is an expected decrease in rule violations and an expected increase in rule following behaviors of your students.

Risks

While there are no foreseeable risks for your students, you may experience some discomfort in learning a new classroom management technique. However, I and/or other

trained graduate students will provide training of each step, feedback, materials necessary, and will be available to answer any questions you may have along the way.

Confidentiality

All interviews, observations, and other information obtained during this study will be kept strictly confidential. Your name, students' names, and other identifying information will not be disclosed to any person not connected with this study. Results from this research project may be shared at professional conferences or published in scholarly journals; however, all identifying information will be removed from publications and/or presentations.

Alternative Procedures

There will be no alternatives offered. However, if any teacher who chooses not to participate will be given the contact information of a USM school psychology liaison who can provide follow-up services to address any classroom management concerns.

Participant's Assurance

This project has been reviewed by the Institutional Review Board, which ensures that research projects involving human subjects follow federal regulations. Any questions or concerns about rights as a research participant should be directed to the Chair of the IRB at 601-266-5997. Participation in this project is completely voluntary, and participants may withdraw from this study at any time without penalty, prejudice, or loss of benefits.

Consent to Participate in Research

Consent is hereby given to participate in this research project. All procedures and/or investigations to be followed and their purpose, including any experimental procedures, were explained to me. Information was given about all benefits, risks, inconveniences, or discomforts that might be expected.

The opportunity to ask questions regarding the research and procedures was given. Participation in the project is completely voluntary, and participants may withdraw at any time without penalty, prejudice, or loss of benefits. All personal information is strictly confidential, and no names will be disclosed. Any new information that develops during the project will be provided if that information may affect the willingness to continue participation in the project.

Questions concerning the research, at any time during or after the project, should be directed to the Principal Investigator (Meleah Ackley; meleah.ackley@usm.edu; 601-270-3071) or Dr. Evan Dart (601-266-4596; evan.dart@usm.edu). This project and this consent form have been reviewed by the Institutional Review Board, which ensures that research projects involving human subjects follows federal regulations. Any questions or concerns about rights as a research participant should be directed to the Chair of the Institutional Review Board, The University of Southern Mississippi, 118 College Drive #5147, Hattiesburg, MS 39406-5147, (601) 266-6820.

Sincerely,

_____	_____
Meleah Ackley, M.S., BCBA	Evan H. Dart, Ph.D., BCBA-D
School-Psychologist-in-Training	Assistant Professor
The University of Southern Mississippi	The University of Southern Mississippi

Please Read, Sign, and Return the Following:

I have read the above documentation and consent to participate in this project. I have had the purpose and procedures of this study explained to me and have had the opportunity to ask questions. I am voluntarily signing this form to participate under the conditions stated. I have also received a copy of this consent. I understand that I will be asked to implement a classroom-based intervention called the Good Behavior Game, and observations will be conducted in the classroom on behavior. In order to take part in this study, I will be required to complete a consultation session and to implement the interventions. In addition, I will be trained on all of the intervention procedures by the primary experimenter or another graduate student. I further understand that all data collected in this study will be confidential and that my name and the students' names will not be associated with any data collected. I understand that I may withdraw my consent for participation at any time without penalty, prejudice, or loss of privilege.

_____	_____
Signature of Teacher	*Date*
_____	_____
Signature of Witness	*Date*

APPENDIX D – Class-wide Direct Behavior Rating (DBR)

Class-wide Rating of Academic Engaged Time

Date: _____ Session: _____

Rater (Teacher): _____ Observer: _____

Start Time: _____ End Time: _____

Instructions: Write an X on a number (0 – 10) that best reflects the percentage of time the class appeared on-task during the duration of the observation. Marks between numbers indicate 5%, 10%, 15%, 25%, and so on. Try to rate the class as an entire group instead of focusing on one specific student.

On-task is defined as those times when the student is attending to the assigned work (e.g., writing, reading aloud, raising a hand, talking to peer about assigned material) or passively attending to assigned work (e.g., looking at the teacher as she speaks to the class, reading assigned material silently). Non-examples include walking around the classroom without permission, calling out, and aimlessly looking around the classroom, silently reading unassigned material.

0	1	2	3	4	5	6	7	8	9	10
0%					50%					100 %
Never					Sometimes					Always

APPENDIX E – Teacher Script

Today we are going to have a competition, and the class will be divided into two teams. Each team has to compete to win a reward. Your team can win by following these rules [read and explain the classroom rules]. If any person on your team does not follow a rule, the team loses a point. If any person does follow a rule, the team gets a point added. As long as you have x [state criterion] or more, then your team will win a prize from the treasure box!

Remember:

- Remind the class of the rules every day.
- When you add a point for a rule following behavior, explain which rule was followed and deliver praise for following the rule.
- When you take away a point for a rule violating behavior, explain which rule was not followed and deliver a reprimand for violating the rule.
- After 20 minutes, announce when game ends.

At the end of the game, allow the winning students to choose a prize.

APPENDIX F – Teacher Reward Menus

Class A Reward Menu	
Reward	**Points**
Expo Marker:	48
Glue Stick:	40
2 AA Batteries:	32
Pen:	28
Pencil:	20

Class B Reward Menu	
Reward	**Points**
Expo Marker:	50
Sticky Note:	42
Tape:	34
Pen:	30
Pencil:	22

Class C Reward Menu	
Reward	**Points**
Expo Marker:	60
2 AA Batteries:	50
Sticky Notes	40
Pen:	30
Pencil:	20

APPENDIX G – Example Data Sheet

Class A VI Schedule_Session16

Observer: _____ Primary/IOA

Date:_____

Praise Schedule	
32	34
20	38
42	34
41	38
37	36
40	42
33	39
36	38
33	35
38	40
31	34
35	37
39	37
33	43
39	
32	

Reprimand Schedule	
47	98
180	89
150	
100	
110	
200	
230	
74	
190	
120	

Total # Points Delivered: _____

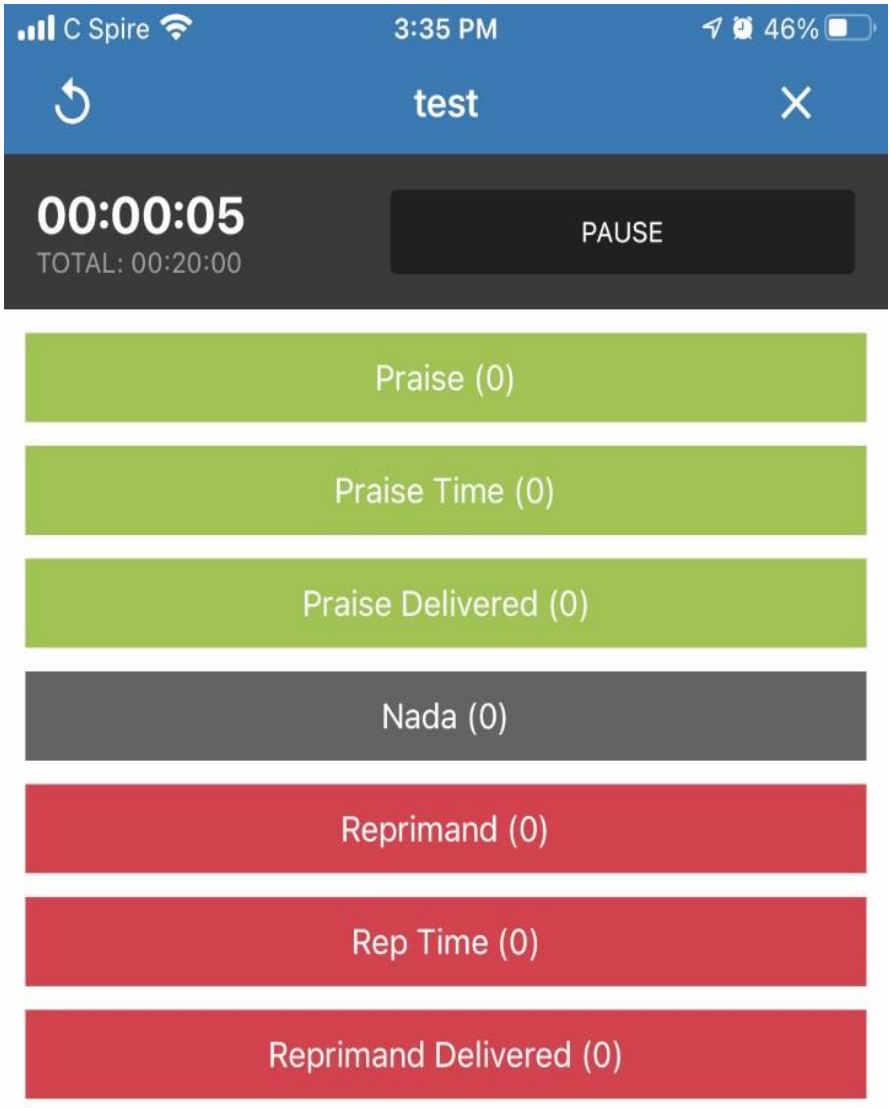
Total # Points Teacher Used: ____

After rewards chosen,

carryover points: ___

(write carryover points on board)

APPENDIX H – COUNTEE



APPENDIX I - GBG Integrity Checklist

TREATMENT INTEGRITY STEPS			
Teacher announces start of the game.	✓	X	N/A
Intervention classroom rules poster is posted.	✓	X	N/A
Teacher states the reward teams are competing for (e.g., prize from treasure box).	✓	X	N/A
Teacher reviews rules with the class.	✓	X	N/A
Criterion level is told to the students and displayed on the board.	✓	X	N/A
Teacher reminds students of how to win the game (e.g., "As long as you have 10 or more points, your team can win.").	✓	X	N/A
Teacher announces when the game has ended.	✓	X	N/A
Teacher determines who won the game.	✓	X	N/A
Teacher allows winning team access to the reward.	✓	X	N/A
Steps completed		/	
Percentage of steps completed			

Taken and adapted from Hunt, B. M. (2012). Using the Good Behavior Game to decrease disruptive behavior while increasing academic engagement with a Headstart population (Unpublished master's thesis). The University of Southern Mississippi, Hattiesburg, MS. and Mitchell, R. R., Tingstrom, D.H., Dufrene, B.A., Ford, W.B., & Sterling, H.E. (2015). The effects of the good behavior game with general education high school students. School Psychology Review, 44 (2), 191-207.

APPENDIX J – Procedural Integrity Checklist

PROCEDURAL INTEGRITY STEPS			
The trainer explains the rules and procedures of the intervention to the teacher.	✓	X	N/A
The trainer reviews the teacher script with the teacher	✓	X	N/A
The trainer role-plays the intervention with the teacher, allowing the teacher to act as a student in the classroom.	✓	X	N/A
The trainer role-plays the intervention with the teacher, allowing the teacher to practice implementing the steps of the game.	✓	X	N/A
The trainer provides appropriate feedback contingent upon teacher mistakes during the role-play implementation session.	✓	X	N/A
The trainer insures the teacher has a full understanding of the intervention components.	✓	X	N/A
The trainer describes the contingency for teachers to earn rewards without mentioning rates of praise or reprimand statements.	✓	X	N/A
Steps completed		/	
Percentage of steps completed			

Taken and adapted from Ford, W.B. (2015). Reducing disruptive behavior in high school: The good behavior game (Master's Thesis). Retrieved from The University of Southern Mississippi: The Aquila Digital Community. and Lambert, A.M., Tingstrom, D.H., Sterling, H.E. Dufrene, B.A. & Lynne, S. (2015). Evaluating the use of tootling for improving upper elementary/middle school students' disruptive and appropriate behavior. *Behavior Modification, 39*(3), 413-430

REFERENCES

Barrish, H.H., Saunders, M., & Wolf, M.W. (1969). Good behavior game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2,* 119-124.

Baum, W.M. (1974a). Choice in free-ranging wild pigeons. *Science, 185,* 78-79.

Baum, W.M.(1974b). On two types of deviation from the matching equation: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22,* 231–242.

Baum, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior, 32,* 269–281.

Beaman, R. & Wheldall, K. (2000). Teachers' use of approval and disapproval in the classroom. *Educational Psychology, 20,* 431-446.

Billington, E., & DiTommaso, N. M. (2003). Demonstrations and applications of the matching law in education. *Journal of Behavioral Education, 12,* 91-104.

Borrero, C.S. W., Vollmer, T.R., Borrero, J.C., Bourret, J.C., Sloman, K. N., Samaha, A. L., & Dallery, J. (2010). Concurrent reinforcement schedules for problem behavior and appropriate behavior: Experimental applications of the matching equation. *Journal of the Experimental Analysis of Behavior, 93,* 455–469.

Borrero, J.C., & Vollmer, T.R. (2002). An application of the matching equation to severe problem behavior. *Journal of Applied Behavior Analysis, 35,* 13-27.

Crowley, M.A., & Donahoe, J. W. (2004). Matching: Its acquisition and generalization. *Journal of the Experimental Analysis of Behavior, 82,* 143-159.

Dart, E. H., Cook, C. R., Collins, T. A., Gresham, F. M., & Chenier, J. S. (2012). Test driving interventions to increase treatment integrity and student outcomes. *School Psychology Review, 41,* 467– 481.

Davison, M., & McCarthy, D. (2016). The matching law: A research review.

Elswick, S., & Casey, L.B. (2011). The good behavior game is no longer just an effective intervention for students: An examination of the reciprocal effects on teacher behaviors. *Beyond Behavior,21* (1), 36-46.

Ford, W.B. (2015). Reducing disruptive behavior in high school: The good behavior game (Master's Thesis). Retrieved from The University of Southern Mississippi: The Aquila Digital Community.

Greguson, L.A. (2008). Differential effects of reinforcement rate and delay on response allocation: A systematic approach utilizing the matching equation. ProQuest.

Herrnstein, R.J. (1970). On the equation of effect. *Journal of the Experimental Analysis of Behavior, 13,* 243-266.

Herrnstein, R.J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4,* 267-272.

Herrnstein, R.J., & Loveland, D. H. (1975). Maximizing and matching on concurrent ratio schedules. *Journal of the Experimental Analysis of Behavior, 24,* 107-116.

Hunt, B. M. (2012). Using the Good Behavior Game to decrease disruptive behavior while increasing academic engagement with a Headstart population (Unpublished master's thesis). The University of Southern Mississippi, Hattiesburg, MS.

Kodak, T., Miltenberger, R.G., & Romaniuk, C. (2003). The effects of differential negative reinforcement of other behavior and non-contingent escape on compliance. *Journal of Applied Behavior Analysis, 36*(3), 379-382.

Lambert, A.M., Tingstrom, D.H., Sterling, H.E. Dufrene, B.A. & Lynne, S. (2015). Evaluating the use of tootling for improving upper elementary/middle school students' disruptive and appropriate behavior. *Behavior Modification, 39*(3), 413-430.

Lannie, A.L., & McCurdy, B.L. (2007). Preventing disruptive behavior in the urban classroom: Effects of the good behavior game on student and teacher behavior. *Education and Treatment of Children, 30*(1), 85-98.

Leflot, G., Lier, P.A.C., Onghena, P., & Colpin, H. (2010). The role of teacher behavior management in the development of disruptive behaviors: An intervention study with the good behavior game. *Journal of Abnormal Child Psychology, 38*, 869-882.

Martens, B.K., Halperin, S., Rummel, J.E., Kilpatrick, D. (1990) Matching theory applied to contingent teacher attention. *Behavioral Assessment, 12*(2), 139-155.

McDowell, J. J. (1988). Matching theory in natural human environments. *The Behavior Analyst, 11*, 95–109.

Mitchell, R. R., Tingstrom, D.H., Dufrene, B.A., Ford, W.B., & Sterling, H.E. (2015). The effects of the good behavior game with general education high school students. *School Psychology Review, 44* (2), 191-207.

Noell, G.H., Witt, J.C., Slider, N.J., Connell, J.E., Gatti, S.L., Williams, K.L., Koenig, J.L., Resetar, J.L. (2005). Treatment implementation following behavioral

consultation in schools: A comparison of three follow-up strategies. *School Psychology Review, 34,* 87-106.

Poling, A., Edwards, T. L., Weeden, M., & Foster, T. M. (2011). The matching equation. *Psychological Record, 61*(2), 313-322.

Reed, D.D. (2009). Using Microsoft Excel 2007 to conduct generalized matching analyses. *Journal of Applied Behavior Analysis, 42,* 867-875.

Reed, D.D., Critchfield, T.S. & Martens, B.K. (2006). The generalized matching equation in elite sport competition: Football play calling as operant choice. *Journal of Applied Behavior Analysis, 39,* 281-297.

Riley-Tillman, T.C., Chafouleas, S.M., Sassu, K.A., Chanese, J.A.M., & Glazer, A.D. (2008). Examining the agreement of Direct Behavior Ratings and Systematic Direct Observation data for on-task and disruptive behavior. *Journal of Positive Behavior Interventions, 10*(2), 136–143.

Riley-Tillman, T.C., Methe, S.A., & Weegar, K. (2009). Examining the use of Direct Behavior Rating on formative assessment of class-wide engagement. Assessment for Effective Intervention, *34*(4), 224–230. doi:10.1177/1534508409333879

Shriver, M.D., & Kramer, J. J. (1997). Application of the generalized matching equation for description of student behavior in the classroom. *Journal of Behavioral Education, 7*(2), 131-149.

Swiezy, N. B., Matson, J. L., & Box, P. (1993). The good behavior game: A token reinforcement system for preschoolers. *Child and Family Behavior Therapy, 14,* 21-32.

Tanol, G., Johnson, L., McComas, J., & Cote, E. (2010). Responding to rule violations or rule following: A comparison of two versions of the Good Behavior Game with kindergarten students. *Journal of School Psychology, 48,* 337-355.

Tingstrom, D. H. (1994). The good behavior game: An investigation of teachers' acceptance. *Psychology in the Schools, 31,* 57-65.

Tingstrom, D., Sterling-Turner, H., & Wilczynski, S. (2006). The good behavior game: 1969-2002. *Behavior Modification, 30,* 225-253.

Vollmer, T. R., & Bourret, J. (2000). An application of the matching equation to evaluate the allocation of two and three-point shots by college basketball players. *Journal of Applied Behavior Analysis, 33,* 137–150.